ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΔΗΜΟΣΙΑΣ ΔΙΟΙΚΗΣΗΣ & ΑΥΤΟΔΙΟΙΚΗΣΗΣ
ΙΝΣΤΙΤΟΥΤΟ ΕΠΙΜΟΡΦΩΣΗΣ

Στατιστική Συμπερασματολογία με Στατιστικά Πακέτα

Παρουσίαση Εκπαιδευτή

Μαθησιακό Αντικείμενο:

Στοιχεία Πιθανοθεωρίας Ι

Εκπαιδευτικοί Στόχοι

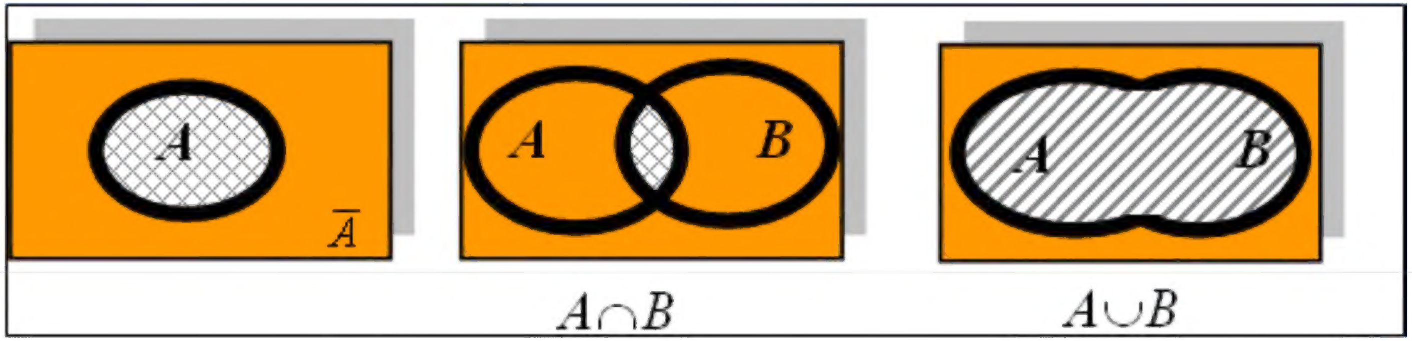
Με την υλοποίηση του μαθησιακού αντικειμένου, ο καθένας από τους συμμετέχοντες θα μπορεί:

- Να κατανοεί τα παρακάτω στοιχειά θεωρίας Πιθανοτήτων: Τυχαίο πείραμα, Ενδεχόμενα, Δειγματικός χώρος.
- Να κατανοεί την έννοια Πιθανότητα και τις βασικές ιδιότητες των πιθανοτήτων.
- Να αντιλαμβάνεται περί τυχαίων μεταβλητών και κατανομών πιθανότητας: έννοιες, κατανομές, παράμετροι και ιδιότητές τους.
- Να αναγνωρίζει και να κατανοεί τις σχέσεις μεταξύ τυχαίων μεταβλητών.

Βασικές έννοιες

- ***Πείραμα* (experiment)** ορίζεται η διαδικασία η οποία μπορεί να επαναλαμβάνεται κάτω από τις ίδιες κάθε φορά συνθήκες και να οδηγεί σε κάποιο από ορισμένα δυνατά αποτελέσματα.
- ***Αποτέλεσμα* (outcome or event)** ενός πειράματος είναι μια παρατήρηση ή μια μέτρησή του. *Για παράδειγμα η ρίψη ενός νομίσματος n φορές στο ίδιο περιβάλλον αποτελεί ένα πείραμα, όπως επίσης οι ιατρικές έρευνες για την επίδραση των φαρμάκων στη θεραπεία των ασθενειών.*
- Όταν τα αποτελέσματα ενός πειράματος δεν είναι γνωστά ή δεν μπορούν να προβλεφθούν με βεβαιότητα τότε μιλάμε για ***πείραμα τύχης***.

Διάγραμμα Venn



Σχήμα 1: Διαγράμματα Venn - σχέσεις ενδεχομένων.

Ενδεχόμενα

- ***Αδύνατο*** ενδεχόμενο είναι εκείνο που δεν περιέχει κανένα αποτέλεσμα του πειράματος τύχης.
- ***Ανεξάρτητα*** ενδεχόμενα είναι εκείνα που η εμφάνιση ενός δεν αποκλείει την εμφάνιση του άλλου.
- ***Αμοιβαία αποκλειόμενα* ή *ξένα*** λέγονται δύο ενδεχόμενα που η έλευση του ενός αποκλείει την εμφάνιση του άλλου.
- ***Βέβαια ενδεχόμενα*** είναι αυτά που θα συμβούν οπωσδήποτε κατά την εκτέλεση του πειράματος τύχης.
- ***Πιθανό ενδεχόμενο*** είναι εκείνο που η εμφάνισή του δεν είναι βέβαιη.

Προσδιορισμός πιθανοτήτων (1)

Οι τρεις ορισμοί-ερμηνείες της απλής ή οριακής ή περιθώριας πιθανότητας (simple or marginal probability) είναι η:

- **Κλασική ή μαθηματική πιθανότητα.**
- **Εμπειρική ή στατιστική πιθανότητα.**
- **Υποκειμενική πιθανότητα.**

Προσδιορισμός πιθανοτήτων (2)

Ο **κλασικός ορισμός** διατυπώθηκε από τον P.S. Laplace και ονομάζεται και *θεωρητική ή μαθηματική ή εκ των προτέρων πιθανότητα.*

Αφορά στα δυνατά και αμοιβαία αποκλειόμενα ενδεχόμενα δειγματικού χώρου με το σημαντικό χαρακτηριστικό ότι είναι όλα ισοπίθανα (equally likely).

$$P(A) = \frac{n(A)}{n(S)}$$

$n(A)$: το πλήθος των ισοπίθανων τρόπων εμφάνισης του ενδεχομένου A (ευνοϊκές περιπτώσεις) και

$n(S)$: το πλήθος όλων των ισοπίθανων ενδεχομένων του δειγματικού χώρου S.

Προσδιορισμός πιθανοτήτων (3)

Ο *R. Von Mises* ορίζει την *εμπειρική ή στατιστική ή εκ των υστέρων Πιθανότητα*, όπου η εμπειρική πιθανότητα του ενδεχομένου *A* εκτιμάται ως το όριο της σχετικής συχνότητας εμφάνισης του *A* όταν ο αριθμός επαναλήψεων, κάτω από τις ίδιες πάντοτε συνθήκες, του τυχαίου πειράματος (τ.π.) είναι πάρα πολύ μεγάλος (τείνει στο άπειρο), συμβολικά ως εξής:

$$P(A) = \lim_{N \to \infty} \frac{n(A)}{N}, \; 0 \le n(A) \le N$$

Όπου N το πλήθος επαναλήψεων του τυχαίου πειράματος.

Ιδιότητες Πιθανοτήτων (1)

$P(A) \geq 0, \forall A$
$P(S) = 1$
$P(A \text{ ή } \Gamma) = P(A) + P(\Gamma)$

Όπου *Α, Γ* ξένα ενδεχόμενα

$0 \leq P(A) \leq 1$
$P(A') = 1 - P(A), \alpha\varphi o\acute{\upsilon} S = A \cup A', A \cap A' = \varnothing \; \kappa\alpha\iota \; P(\varnothing) = 0 \Rightarrow$ $P(S) = P(A) + P(A') = 1 \Rightarrow P(A') = 1 - P(A)$

Ιδιότητες Πιθανοτήτων (2)

Ο ορισμός της ***απλής ή οριακής πιθανότητας*** ως άθροισμα συνδυασμένων:

$$P(A) = P(A\,\kappa\alpha\iota\,B_1) + P(A\,\kappa\alpha\iota\,B_2) + \ldots + P(A\,\kappa\alpha\iota\,B_N)$$

Όπου B_1, B_2, ..., B_N είναι N αμοιβαία αποκλειόμενα ενδεχόμενα που εξαντλούν το δειγματικό χώρο.

Γενικός προσθετικός κανόνας των πιθανοτήτων:

$$P(A\,\acute{\eta}\,B) = P(A) + P(B) - P(A\,\kappa\alpha\iota\,B) \quad \text{ή}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ειδικός κανόνας της πρόσθεσης πιθανοτήτων:

$$P(A\,\acute{\eta}\,B) = P(A) + P(B) \quad \text{ή}$$

$$P(A \cup B) = P(A) + P(B)$$

Ιδιότητες Πιθανοτήτων (3)

Γενικός κανόνας του πολλαπλασιασμού των πιθανοτήτων:

$$P(A \; και \; B) = P(A|B)P(B) \quad ή \quad P(A \cap B) = P(A|B)P(B)$$

Νόμος του πολλαπλασιασμού για ανεξάρτητα ενδεχόμενα:

$$P(A \; και \; B) = P(A)P(B) \quad ή \quad P(A \cap B) = P(A)P(B)$$

Ορισμός της οριακής ή περιθώριας πιθανότητας:

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_N)P(B_N)$$

Όπου B_1, B_2, ..., B_N είναι N αμοιβαία αποκλειόμενα ενδεχόμενα που εξαντλούν το δειγματικό χώρο.

Τυχαίες μεταβλητές και συναρτήσεις πιθανότητας (1)

Η συνάρτηση που ως πεδίο ορισμού έχει τις τιμές x_i, i=1,2,3,…,n, π.χ. διακριτής τ.μ. X, και πεδίο τιμών τις αντίστοιχες πιθανότητές τους $f(x_i)$ ονομάζεται ***συνάρτηση πιθανότητας*** *της* ***ασυνεχούς*** τ.μ. X και γράφεται ως εξής:

$$f(x_i) = P_x = P(x_i) = P(X = x_i), \quad i = 1, 2, \ldots, n$$

αν ικανοποιούνται οι συνθήκες

$$f(x_i) \geq 0 \quad \kappa\alpha\iota \quad \sum_i f(x_i) = 1$$

Τυχαίες μεταβλητές και συναρτήσεις πιθανότητας (2)

Η αντίστοιχη έννοια για τις **συνεχείς** τ.μ. ονομάζεται ***συνάρτηση πυκνότητας πιθανότητας*** και ορίζεται ως εξής:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)dx, \text{ με } \alpha < \beta$$

αν ικανοποιούνται οι συνθήκες

$$f(x) \geq 0 \quad , -\infty < x < +\infty \quad \text{και} \quad \int_{-\infty}^{+\infty} f(x)dx = 1$$

Τυχαίες μεταβλητές και συναρτήσεις πιθανότητας (3)

Συνάρτηση Κατανομής ή αθροιστική συνάρτηση κατανομής:

$$F(x) = P(-\infty < X \leq x) = \begin{cases} \sum_{X \leq x} f(x), & \alpha\nu \text{ X } \delta\iota\alpha\kappa\rho\iota\tau\acute{\eta} \\ \int_{-\infty}^{x} f(x)dx, & \alpha\nu \text{ X } \sigma\upsilon\nu\varepsilon\chi\acute{\eta}\varsigma \end{cases}$$

Παράμετροι τυχαίων μεταβλητών (1)

Ο ***μέσος μιας τυχαίας μεταβλητής*** X με συνάρτηση πυκνότητας πιθανότητας f(x) είναι η αναμενόμενη τιμή της (expected value) **E(X)**, η οποία ορίζεται ως ο μέσος σταθμικός των τιμών της X με στάθμιση τις αντίστοιχες πιθανότητές τους.

$$E(X) = \begin{cases} \sum_i x_i f(x_i) = \mu, & \alpha\nu\ X\ \delta\iota\alpha\kappa\rho\iota\tau\acute{\eta} \\ \int_{-\infty}^{+\infty} x f(x) dx = \mu, & \alpha\nu\ X\ \sigma\upsilon\nu\varepsilon\chi\acute{\eta}\varsigma \end{cases}$$

Αν X τυχαία μεταβλητή με μέσο **E(X)=μ**, τότε η προσδοκώμενη τιμή των τετραγωνικών αποκλίσεών της από το μέσο, $[X-E(X)]^2$, ονομάζεται ***διακύμανση* της X** και ορίζεται ως εξής:

$$\sigma^2 = \mathrm{var}(X) = E[X - E(X)]^2 = \begin{cases} \sum_i (x_i - \mu)^2 f(x_i) = \mu, & \alpha\nu\ X\ \delta\iota\alpha\kappa\rho\iota\tau\acute{\eta} \\ \int_{-\infty}^{+\infty} (x_i - \mu)^2 f(x) dx = \mu, & \alpha\nu\ X\ \sigma\upsilon\nu\varepsilon\chi\acute{\eta}\varsigma \end{cases}$$

Παράμετροι τυχαίων μεταβλητών (2)

Μια άλλη σημαντική παράμετρος για τη μέτρηση της συμμεταβολής *δύο* τυχαίων μεταβλητών είναι η ***συνδιακύμανσή* τους**.
Η συνδιακύμανση μετράει τη μορφή (θετική ή αρνητική) και το βαθμό ή την ένταση της *γραμμικής* συσχέτισης (συνάφειας) τους στις απόλυτες μονάδες μέτρησής τους και συμβολίζεται:

$$Cov(X,Y) = E\{[x_i - E(X)][y_i - E(Y)]\} = E[(X-\mu_x)(Y-\mu_y)]$$
$$= \sum_i \sum_j [x_i - E(X)][y_j - E(Y)] f(x_i y_j)$$
$$= \sum_i \sum_j x_i y_j f(x_i y_j) - \mu_x \mu_y$$

όπου $\mu_x = \sum_i x_i f(x_i)$ και $\mu_y = \sum_j y_j f(y_j)$.

Παράμετροι τυχαίων μεταβλητών (3)

Όταν θέλουμε να μετρήσουμε τη συμμεταβολή των X, Y χρησιμοποιώντας στατιστικό απαλλαγμένο των μονάδων μέτρησής τους, χρησιμοποιούμε το *συντελεστή γραμμικής συσχέτισης* του Pearson που υπολογίζεται από τη σχέση:

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\sum_i \sum_j [x_i - E(X)][y_j - E(Y)] f(x_i y_j)}{\sqrt{\sum_{i=1}^{n} [x_i - E(X)] f(x_i)} \sqrt{\sum_{j=1}^{k} [y_j - E(Y)] f(y_j)}}$$